Converging Tides Lift All Boats: Consensus in Evaluation Criteria

Boosts Investments in Firms in an Emerging Technology Field

(online appendix can be accessed at

https://www.dropbox.com/s/lx53a2t2ftiwj9v/OS_specialconference20200504appendix.pdf?dl=0)

Abstract

While previous studies show that the emergence of evaluation criteria for a new technology improves the life chances of well-performing firms, we theorize that consensus in such criteria among technology experts increases investments to *all* firms in the new sector, and provide supportive evidence from an experiment with 80 Chinese investors (Study 1). We further explore this result in a second experiment with 412 U.S. participants (Study 2), showing that evaluation criteria consensus among technology experts both increases investors' propensity to view a firm as technologically competent and to expect that other investors will favor investing in the firm. Analyses of archival data on investment in artificial intelligence technology firms by U.S. and Chinese investors (Study 3a and Study 3b, respectively) also lend support to our arguments. By exploring the social-cognitive processes that link evaluation criteria consensus among experts to investors' assessment of firms in emerging technology fields, this paper advances scholarly understanding of the micro-foundations of the institutionalization processes in new market sectors.

Key words: evaluation criteria, evaluation criteria consensus, investment, emerging technology, micro-foundations of institutional theory, experimental methods

Converging Tides Lift All Boats: Consensus in Evaluation Criteria

Boosts Investments in Firms in an Emerging Technology Field

For investors, firms in emerging technology fields are thorny roses (Aldrich and Fiol 1994; Tushman and Anderson 1986). On the one hand, making an investment in a new firm whose technology may create market-changing opportunities can yield extraordinary returns. On the other hand, evaluating the core technology of the firm may be a challenge (Suchman 1995; Tushman 1992). Difficulties in evaluating firm technologies in the nascent stage of a sector undermine investors' confidence and willingness to provide resources that are critical to the sector's firms (Suchman 1995) and the field as a whole. Past research has examined field-level evaluation institutions, such as certification agencies (Graffin and Ward 2010; Lanahan and Armanios 2018; Sine et al. 2007) and technology competitions (Rao 1994), finding that certificates from these agencies and participation in these competitions increases the survival rates of firms (Goldfarb et al. 2018; Rao 1994; Sine et al. 2007). A conclusion is that the development of evaluation criteria in a nascent sector reduces the uncertainty resource providers face (Graffin and Ward 2010) and thus benefit some firms.

However, existing research on evaluation has overlooked the consequences of having consensus (as opposed to disagreements) in evaluation criteria for new technologies. We define consensus in evaluation criteria as the extent to which technology experts in a nascent sector use similar criteria (Hsu et al. 2012) to compare different technology solutions. Technology experts are academic and industry scientists who conduct research on the technology, and thereby are recognized as having domain-specific knowledge (Dew et al. 2009; Kynn 2008; Morgan 2014). There can be varying levels of consensus because multiple evaluation criteria often exist in nascent fields (Chatterji et al. 2016; Lee et al. 2010; Yue et al. 2013). For example, in the early automotive industry, various criteria, such as speed, durability, and fuel efficiency, were used to assess firms' technological performance in different technology competitions (Rao 1994). Nevertheless, existing research has not theorized nor empirically examined the consequences of having consensus (as opposed to disagreements) in evaluation criteria. This paper addresses this gap, focusing on the effect of evaluation criteria consensus on investment in firms that rely on emerging technologies.

We argue that field-level consensus among technology experts on how to evaluate a new technology is essential for investors' decision making. One key problem facing investors in emerging technology fields is a lack of proper criteria for assessing the relative performance of alternative forms of the new technology. The evaluation criteria developed by experts often become convenient and legitimate tools that investors use to evaluate firms' technology (Morgan 2014). According to a research scientist at a large Internet company who organized many early criteria-setting technology competitions in artificial intelligence, "investors were often in the audience of our competitions". Consensus in evaluation criteria makes it easier for firms using the technology to secure investment for two reasons. First, it serves as a signal that the new technology is viable, and investors are apt to generalize this inference to all firms that utilize this technology. Therefore, consensus on evaluation criteria for a technology enhances investors' positive perceptions of the technological competence of firms in the sector. By perceptions of technological competence, we mean investors' assessment of the capability and reliability of the firm's technological solutions. Second, consensus on evaluation criteria increases investors' confidence that other people will also hold a positive view of the focal firm, making it easier to justify their decisions to their stakeholders. Therefore, the investors are likely to perceive others as willing to invest in these firms like themselves. Both perceptions should encourage investors to invest in these firms.

We tested our theoretical propositions in two experiments and two archival studies. In our first experiment, with 80 Chinese private equity investors as our participants, we find that expert consensus on criteria for evaluating a new technology increases investors' willingness to invest in a firm in the sector. In a second experiment with 412 U.S. participants, we show that evaluation criteria consensus boosts investment because participants are more likely to perceive a firm as technologically competent and others as evaluating the firm positively. To establish the external validity of our findings, our third and fourth studies use longitudinal archival data on investment in artificial intelligence (hereafter referred to as AI) from the U.S. and China. To gauge the degree of evaluation criteria consensus, we collected data from competitions held at major AI conferences that were designed to compare the performance of alternative forms of particular AI technologies and examined the variation in performance metrics across different competitions. Both studies demonstrate that evaluation criteria consensus for a technology increases the likelihood of a firm receiving investment.

This study contributes to organizational studies of emerging technology fields in three ways. First and foremost, we contribute to the literature on emerging market sectors (e.g. Grodal 2007; Hiatt et al. 2009; Santos and Eisenhardt 2009) by highlighting the positive impact of field-level consensus in evaluation criteria on new venture resource acquisition. Extending existing research which emphasizes field-level institutions such as agreed-upon market architecture (Ozcan and Santos 2015) and producer identities (Georgallis et al. 2018), this study is the first, to our knowledge, to explicitly theorize and measure *consensus* in experts' evaluation criteria and its implications for investment. Moreover, we extend existing research on the impact of evaluation criteria in emerging market sectors. In contrast to studies which show that only high-performing firms (Rao 1994) or ones that have been certified (Sine et al. 2007) benefit from the establishment of technology evaluation criteria

(Rao 1994; Sine et al. 2007), we show that any firm in the field is likely to benefit from field-level consensus in evaluation criteria. In other words, we establish "a converging tide lifts all boats" effect.

Second, by showing how consensus among experts affects individual investors' perceptions and decisions, this study is among the first to provide micro-level evidence of institutionalization processes in emerging technology fields. Existing research on legitimating institutions in emerging fields (e.g., Grodal 2018; Lee et al. 2018; Lounsbury and Glynn 2001) primarily relies on macro-historical data (for a recent exception, see Zunino et al. 2019), which makes it hard to document or draw causal inferences about micro-foundational processes (Bitektine et al. 2018). Moreover, while past studies (Rao 1994; Sine et al. 2007) indicate that the establishment of evaluation criteria affects firm survival, they did not directly examine the underlying processes. Our two experiments not only establish a causal relationship between consensus in the evaluation criteria and investment outcomes, but also highlight two mediators of this relationship—investors' perceptions of (a) a firm's technological competence and (b) others' evaluations. Our findings also extend research on investment in emerging technology fields (Aggarwal et al. 2015; Baum and Silverman 2004; Fisher et al. 2017; Hsu 2007) by showing that perceptions of technological competence, a key factor which predicts investments (Aggarwal et al. 2015; Baum and Silverman 2004), is shaped by cognitive institutions in the field.

Third, we contribute to the literature on technology evolution and industry emergence (Grodal et al. 2015; Moeen and Agarwal 2017; Suárez and Utterback 1995). While prior literature has shown that technology standards (Tassey 2000; Vakili 2016) and designs (Anderson and Tushman 1990; Suárez and Utterback 1995) reduce uncertainty in nascent fields by coordinating transactions among market actors. We discover what reduces this uncertainty before the existence of agreed-upon technology standards and designs. We focus on consensus in evaluation criteria, an understudied cognitive institution (Goldfarb et al. 2018; Lanahan and Armanios 2018; Sine et al. 2007). Contributing to the growing literature on cognitive institutions in emerging industries (e.g., Bingham and Kahl 2013; Grodal 2007; Kaplan and Tripsas 2008; Lounsbury et al. 2003), we establish that evaluation criteria consensus boosts firm investment.

Theory and Hypotheses

Evaluation of Firms in Emerging Technology Fields

A crucial factor that investors consider when deciding whether to invest in a technology-based firm is its technological competence (Aggarwal et al. 2015; Baum and Silverman 2004). However, making such assessment is inherently difficult in a nascent technology field, because of two types of uncertainty, which Graffin and Ward (2010) refer to as *technical uncertainty* and *performance standard uncertainty*. The first arises from a lack of observable data and well-defined objective metrics for gauging the quality of performance. The latter is created by a lack of collective agreement on an acceptable performance level. In this paper, we define *evaluation criteria* for a technology as a scheme that (a) defines the observable features/functions on which firms' technological performances should be measured (Durand and Kremp 2016; Durand and Paolella 2013) and the objective metrics for measuring these features/functions; and (b) specifies an acceptable performance level, either by providing an absolute threshold or a relative comparison. By definition, evaluation criteria resolves both the technical uncertainty and performance standard uncertainty (Graffin and Ward 2010) in assessing firm technological competence in a nascent sector.

Past literature has documented that the development of evaluation criteria in a new sector can assure investors of some firms' technological competence. For example, in the early days of the automotive industry, technology competitions ranked the performances of different automotive firms on given criteria – speed, endurance, operating costs and so forth. In a study of these competitions, Rao (1994) showed that firms that won more competitions attracted more resources and were more likely to survive. A follow-up study by Goldfarb et al. (2018) found that it was not just the winning firms that were more likely to survive—those that placed second or third also enjoyed enhanced survival chances. Furthermore, these authors found that mere participation in the competitions could be as beneficial as winning. This suggests that simply being subjected to an evaluation may affect audiences' perceptions of firms' viability. Similarly, several studies showed that the establishment of agencies that provided certifications to firms that met their criteria also enhanced the survival rates of those firms, regardless of the content of the criteria (Lee et al. 2017; Sine et al. 2007).

The existing literature, however, leaves several important questions unanswered, which motivate this study. First, while existing studies suggest that the presence of evaluation criteria in a new sector positively affects firm survival, they overlook the consequences of coexisting or even competing evaluation criteria (Chatterji et al. 2016; Lanahan and Armanios 2018; Lee et al. 2010). The quick growth of a nascent sector often attracts technology experts from diverse backgrounds with, different interpretations of the defining features/functions of the emerging technology (Grodal 2018; Lee et al. 2017; Wry et al. 2011), and thus, multiple, competing evaluation criteria. In consequence, experts in an emerging field could lack *consensus* in their evaluation criteria. However, little prior work has theorized evaluation criteria *consensus* (for a relevant theoretical discussion, see Lanahan and Armanios 2018) or examined its effects on key resource providers, such as investors. We address this gap by examining how consensus (versus disagreements) in evaluation criteria influences investors' confidence in the technological viability of a new sector, and hence, in its firms.

Second, although existing studies theorize that the establishment of evaluation criteria benefit firms by inducing positive evaluations from key resource providers, little work has empirically examined how evaluation criteria affect resource providers' evaluation process. Oftentimes, resource providers' evaluations were inferred from the firms' survival or performance outcomes (Goldfarb et al. 2018; Rao 1994; Sine et al. 2007). Thus, our field needs direct tests of the processes through which evaluation criteria drives resource providers' assessments of investment opportunities. In particular, Prior studies have not focused on investors, who provide crucial financial and network resources for technology firms (Florin et al. 2003; Hsu 2007). Addressing these gaps, in this paper, we directly measure/manipulate the level of evaluation criteria consensus in an emerging field and examine its impact on investors' perceptions of and investment in its firms.

Evaluation Criteria Consensus and Investment to Firms

We conceptualize consensus in evaluation criteria for an emerging technology as the extent to which technology experts—academic and industry scientists with domain-specific knowledge on the technology—use similar criteria (Hsu et al. 2012) to compare different technology solutions[i]. Experts play an important role in providing criteria for evaluation, especially in the sectors with high uncertainties (Biglaiser 1993; Biglaiser and Friedman 1994; Zuckerman 1999). However, experts' judgments may differ, depending on their heuristics (Kynn 2008; Morgan 2014), institutional and organizational affiliations (Hoffman 1999), political interests, and personal incentives (Campbell 1998), resulting in a lack of consensus on the evaluation criteria for an emerging technology.

Existing studies of emerging market sectors have examined the consequences of having various types of field-level consensus. For example, Ozcan and Santos (2015) showed that the lack of consensus on market architecture, or transactional arrangements, among potential exchange partners from different industries prevented the formation of a global market for mobile payments. Similarly, Georgallis et al. (2018) found that country-level convergence in producers' identities in the new solar photovoltaic industry increased government policy support to the producers (for similar findings, see McKendrick et al. 2003; Negro et al. 2010, 2011). Our study builds on these insights and extends them. First, we focus on consensus among experts in the criteria they use to evaluate a new technology, a crucial cognitive institution that has not been examined to date. More importantly, our study provides a micro-foundation to the historical studies, which primarily rely on historical data (Gurses and Ozcan 2015; Ozcan and Santos 2015; Yue et al. 2013), by explicating the social-cognitive processes through which experts' consensus on evaluation criteria affects resource providers' (i.e. investors') reactions to firms. Relatedly, existing studies often use proxies to measure field-level consensus (e.g. Georgallis et al. 2018). We advance these studies methodologically by manipulating consensus on evaluation criteria in our experiments and quantitatively measure differences in consensus over time and across different emerging technology fields in our archival studies. Our study thus presents a more accurate characterization of the impact of field-level consensus among technology experts on resource providers' decision-making

As elaborated below, we argue that consensus among technology experts on evaluation criteria affects investment in new technology firms through two mechanisms: fostering investors' positive perceptions of firms' technological competence and propensity to expect others to also hold positive perceptions of the firms. Figure 1 depicts our framework.

Insert Figure 1 about here

Perceived technological competence as a mediator. The decision to invest in a firm specializing in an emerging technology depends on an assessment of the firm's technological competence (Aggarwal et al. 2015; Baum and Silverman 2004). However, in a nascent sector, data on the technological competence of a firm is necessarily limited (Plummer et al. 2016; Pollock and Gulati 2007). As a result, investors rely on signals about the firm and its environment (MacMillan et al. 1985; MacMillan et al. 1987) such as compositions of the management team, financial status, and target market (Mason and Stark 2004). For example, venture capital investors base their assessment of firms on the perceived viability of the entire technological field in which the firm is situated (Mason and Stark 2004; Tyebjee and Bruno 1984). Consensus in evaluation criteria induces a sense that the entire technology field is viable and the specific firm's technology solution is reliable. Such a general judgment extends to all firms in the field. Thus, evaluation criteria consensus should foster positive perceptions of the technological competence of any firm in an emerging field. Subsequently, perceived technological competence of a firm should increase investment in it (Aggarwal et al. 2015; Baum and Silverman 2004). Taken together, we propose that perceived technological competence mediates the positive relationship between evaluation criteria consensus and investment in a firm.

It should be noted that our hypothesized mediator is investors' <u>perceptions</u> of firms' technological competence rather than true technological competence. In a nascent field, true technological competence is hard, if not impossible, to identify. This is because investors lack accurate quality signals and may be overloaded with noisy signals[ii] (Plummer et al. 2016; Pollock and Gulati 2007). Moreover, signals about technological competence influence investment via investor's perceptions of technological competence. Eventually it is investor's own perceptions that drive their investment decisions. Thus, we chose the more proximal predictor of investor decision—investor's perceptions of technological competence—as our mediator. Nevertheless, in response to an interesting question raised by a reviewer of this paper, we investigate the effect of a firm's standing on more objective indicators of firms' technological competence in an additional experiment. Please see online Appendix G for details.

Prediction of others' evaluation as a mediator. Evaluation criteria consensus also helps to resolve investors' uncertainty about how their colleagues and other investors would assess the value of a firm (Khaire and Wadhwani 2010). Resolving other-oriented uncertainty (Correll et al. 2017) is particularly important for investors because they face pressures to communicate and justify their decisions (Tetlock et al. 1989) to their colleagues and supervisors (Jensen 2006). More importantly, they are concerned about how to sell the company to other investors—either to syndicate partners in the short run, or as a means of investment exit in the long run (Cumming and binti Johan 2008). Investors would feel more comfortable with their investment in a firm when they predict other key stakeholders to have similar opinions. Anticipating that evaluation criteria consensus would lead other investors to draw the same positive inference about the firm, investors feel more comfortable with investing in the firm.

Hypothesis 1: Expert evaluation criteria consensus in an emerging technology field increases investment in firms in the field.

Hypothesis 2: Investors' perception of firms' technological competence mediates the positive relationship between expert evaluation criteria consensus and investment.

Hypothesis 3: Investors' prediction of others' positive evaluations mediates the positive relationship between expert evaluation criteria consensus and investment.

Scope conditions. As indicated by our title, we state our theoretical claims (that evaluation criteria consensus boosts investment) to be within the scope of emerging technology fields. We expect evaluation criteria consensus to boost investment in an emerging technology field because evaluative uncertainty is high in such fields. Critically, it is under uncertainty that investors rely on consensus in evaluation criteria to infer the firm's technological competence and others' evaluations of the firm, which ultimately shape their investment decisions. Therefore, uncertainty in evaluating the focal technology of a firm constitutes a scope condition for our predictions. Evaluative uncertainty varies depending on how nascent a firm's technology and industry are. We reason that as the nascence level of (and thus uncertainty regarding) a technology/industry varies, our hypothesized effects may change. Thus, we conducted an additional experiment to test whether the positive effect of evaluation criteria consensus on investment holds across technologies and industries with different nascence levels. Please see Appendix H for details.

Overview of Studies

To test our hypotheses, we conducted four studies across two different methodologies (experiments and archival studies) and two national contexts (the U.S. and China). Across the four studies, we examine how consensus on the evaluation criteria for an emerging technology affect investment to firms that draws on the technology. In Study 1, we conducted an experiment with 80 experienced Chinese investors to establish the positive causal relationship between evaluation criteria consensus and investment intentions (Hypothesis 1). Study 2 aimed to replicate the findings in Study 1 with a sample of 412 full-time employees in the U.S. and extend our findings by testing the mediating effects of investors' perceptions of firm technological competence (Hypothesis 2) and predictions of others' positive evaluations (Hypothesis 3). In Study 3a and 3b, we utilize archival data on AI firms in both the U.S. and China to establish external validity of our results. Data for all studies can be accessed at https://osf.io/f7bvw/?view_only=72bddfbaa2e64c23a8e96d8d48d036ea.

Study 1

Study 1 aims to test Hypothesis 1, positing a causal relationship between evaluation criteria consensus and investments. The design and analysis plans for this study were pre-registered and can be accessed at https://osf.io/k8mbe/?view_only=23eba785b4d44e3386e61957d4563637.

Sample

A total of 80 Chinese investors in the private equity market participated in our survey experiment. Participants were recruited from an online investor discussion group affiliated with the alumni association of a top research university in China. Participants in this discussion group are verified alumni who invest in the private equity market and use this discussion group to network and exchange information. Our survey was set up on Qualtrics (www.qualtrics.com), a widely used online survey platform, which accounts were provided to the researchers by our university. The survey link was posted by an active member of the discussion group. Participation was voluntary, anonymous, and without any monetary or other type of reward. As pre-registered, we constrained the eligible participants to only people with investment experience in the private equity market. We excluded 17 participants who did not meet this criterion from our analysis. 153 of all 443 members of the discussion group completed the study (response rate = 34.53%), among which 136 had investment experience (in years, $M = 3.93$ and $SD = 2.74$). Our study included some simple attention checks. As pre-registered and consistent with prior research (e.g., Hildreth, Gino, & Bazerman, 2016), participants who failed attention checks were removed (Appendix C provides details on the attention checks). A total of 80 (58.82%) participants passed the attention checks and were included in our analysis.[iii] 15 were female (18.75%) and 64 male. $M_{age} = 33.28$ and $SD_{age} = 5.24$. None of the participants were aware of our research purposes.

Design and Procedures

The task resembles the first-round screening process at venture capital firms (MacMillan et al. 1987; Tyebjee and Bruno 1984). Appendix A provides details on the task. Each participant was instructed to imagine him/herself as an investor who works in a team at a venture fund. Part of his/her job was to read information about different technologies and firms, beginning with a brief introduction about a start-up firm that makes intelligent cameras, the firm's product, and the entrepreneur. The information that we presented to the participants was adapted from a start-up on Indiegogo, a website that crowd-funds innovative products and features entrepreneurial campaigns.

Participants also read a description of the firm's core technology: image classification algorithms, and a paragraph about evaluation criteria for the technology: "Prominent academic and industry researchers at Stanford, MIT, Google, and Uber all have worked on assessing the algorithms. They have proposed various evaluation metrics for assessing the accuracy of image classification algorithms. The metrics include average precision, ROC-AUC, top-5 error, and others."

Participants were then randomly assigned into one of two conditions in a one-factor two-level (evaluation criteria consensus: high versus low) between-subject experimental design. The random assignment was automatically conducted by the Qualtrics platform. In the high (/low) *evaluation criteria consensus* condition, participants read: "(However,) The academic and industry experts have agreed (do not agree) on the appropriate metrics. In other words, there is great (no) consensus about what metrics should be used for evaluating image classification algorithms."

Lastly, participants completed a survey with measurements for the dependent variables, investment intention and investment amount, the attention checks, the manipulation check, and their demographic information. Appendix B provides details on the measures.

Measures

Investment intention. Three items adapted from Larsen and Newton-Smith (2001) measured intention to invest in the firm. Participants rated the extent to which they agreed or disagreed with statements such as, “I am willing to invest money in this firm.” For all Likert scales in this paper, unless otherwise noticed, 1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neither agree nor disagree, 5 = somewhat agree, 6 = agree, and 7 = strongly agree. $\alpha = 0.95$.

Investment amount. Participants were also asked, “How much of the ￥2,000,000 will you invest in this firm?” The number reported was the measure for investment amount (Kanze et al. 2018).

Manipulation check of evaluation criteria consensus. Three items measured participants’ perceptions of evaluation criteria consensus, asking them to rate the extent to which they agreed or disagreed with statements such as, “There is consensus about the proper metrics to evaluate image classification algorithms.” $\alpha = 0.89$.

Results and Discussion

Manipulation check. A t-test shows a significant difference in perceived evaluation criteria consensus between the high and the low conditions (Difference = 1.97, 95% CI = [1.35, 2.59], d =1.16, t (78) = 6.33, $p < 0.001$; $M_{\textit{high-consensus}} = 4.68$ and $M_{\textit{low-consensus}} = 2.71$), showing that our manipulation worked as intended.

Hypotheses testing. Supporting our Hypothesis 1, as shown in Figure 2, a t-test reveals a significant difference in investment intention between the high and low evaluation criteria consensus conditions (Difference = 0.69, 95% CI [0.10, 1.29], $d = 0.50$, t (78) = 2.32, $p = 0.02$; $M_{\textit{high-consensus}} = 3.88$ and $M_{\textit{low-consensus}} = 3.19$). Similarly, a t-test also suggests a significant difference in investment amount between the high and low evaluation criteria consensus conditions (Difference = 225,000.00, 95% CI [71,900.00, 378,040.00], d =0.63, t (80) = 2.94, p = 0.004; $M_{high-consensus}$ = 437,500.00 and $M_{low-consensus}$ = 212,500.00). Study 1 shows that evaluation criteria consensus increases both investors' investment intention and their investment amount. Results from this experiment provide preliminary causal evidence that evaluation criteria consensus increases investments.

Insert Figure 2 about here

Study 2

The aim of Study 2 is twofold. First, we aim to replicate the positive effect of evaluation criteria consensus on investment (Hypothesis 1) in another social context (the U.S.) and with a larger and more diverse sample. Second, we examine two mechanisms that produce this effect: investors' perceptions of a firm's technological competence (Hypotheses 2) and their perceptions of others' evaluations of the firm (Hypotheses 3). The design and analysis plans were pre-registered and can be accessed at https://osf.io/wnrd7/?view_only=5c458adf013c4b108e637f42f53c0a7d.

Sample

A total of 423 full-time employees with managerial experience in the U.S., drawn from an online platform (https://www.prolific.co/), participated in our experiment. We restricted participants to people who have managerial experience because research shows that managerial experience exposes people to entrepreneurship-relevant information and resources (Tonoyan et al. 2019) and thus prepares them for work related to equity investment. As typical for participants on this and other similar platforms, each participant received a $0.60 reward. We included the same attention checks as those in Study 1 and excluded participants that failed any of the checks. A total of 412 (97.40%) participants passed the attention checks and were included in our analysis. Of these, 178 were female, 226 male, and 8 other. $M_{age} = 37.14$ and $SD_{age} = 10.62$.

Design and Procedures

Same as Study 1, this study was set up on Qualtrics. Participants were randomly assigned into one of two conditions in a one-factor two-level (evaluation criteria consensus: high versus low) between-subject experimental design. The scenario and manipulations for evaluation criteria consensus were the same as those in Study 1. Similarly, participants completed a survey that measured the investment intention and amount, attention checks, a manipulation check, and participants' demographic information. In addition to the measures in the previous study, participants completed the scales of the two proposed mediators, perceived technological competence of the firm and prediction of others' evaluation of the firm. Appendix B provides details on the measures.

Measures

Investment intention. The items were the same as in Study 1. $\alpha = 0.96$.

Investment amount. The question was the same as in Study 1.

Perceived technology competence. Participants completed a 3-item scale adapted from Tyebjee and Bruno (1984). They rated the extent to which they agreed or disagreed with statements such as, "This firm has demonstrated extraordinary technology competence." $\alpha = 0.88$.

Prediction of others' evaluation. Participants completed a 3-item scaleα to rate the extent to which they agreed or disagreed with statements such as, "Other people in my field (e.g., my boss, colleagues, and other investors) will be willing to invest money in this firm." $\alpha = 0.95$.

Manipulation check. The items were the same as in Study 1. $\alpha = 0.96$.

Results and Discussion

Manipulation check. A t-test shows a significant difference in perceived evaluation criteria consensus between the high and the low conditions (Difference = 3.93, 95% CI [3.69, 4.17], $d = 1.68$, $t\,(410) = 31.68$, $p < 0.001$; $M_{\textit{high-consensus}} = 5.93$ and $M_{\textit{low-consensus}} = 2.00$), showing that our manipulation worked as intended.

Main effect. Supporting Hypothesis 1 and replicating the results in Study 1, a t-test reveals a significant difference between the high consensus condition and the low condition in investment intention (Difference = 0.60, 95% CI = [0.31, 0.88], $d = 0.40$, $t\,(410) = 4.14$, $p < 0.001$; $M_{\textit{high-consensus}} = 4.76$ and $M_{\textit{low-consensus}} = 4.16$) and in investment amount (Difference = 20,550.00, 95% CI = [11,684.00, 28,521.00], $d = 0.45$, $t\,(410) = 4.69$, $p < 0.001$; $M_{\textit{high-consensus}} = 64{,}490.00$ and $M_{\textit{low-consensus}} = 44{,}390.00$).

Mediation analysis. To test perceived technological competence and prediction of others' positive evaluation as the mediators, we conducted mediation analysis with 5,000 bootstrap samples (Hayes 2017; Preacher and Hayes 2004). Supporting Hypothesis 2, perceived technological competence mediates the positive indirect effects of evaluation criteria consensus on both investment intention (indirect effect $\hat{a}\hat{b} = 0.63$, 95% CI [0.41, 0.84], $p < 0.001$, standardized indirect effect $\hat{a}\hat{b}_{cs} = 0.21$) and investment amount (indirect effect $\hat{a}\hat{b} = 15{,}129.00$, 95% CI [9,770.00, 20,510.00], $p < 0.001$, standardized indirect effect $\hat{a}\hat{b}_{cs} = 0.17$). In support of Hypothesis 3, predictions of others' positive evaluation also mediates the positive effects of evaluation criteria consensus on both investment intention ($\hat{a}\hat{b} = 0.37$, 95% CI [0.17, 0.57], $p < 0.001$, $\hat{a}\hat{b}_{cs} = 0.12$) and investment amount ($\hat{a}\hat{b} = 9{,}170.00$, 95% CI [4,170, 14,270], $p < 0.001$, $\hat{a}\hat{b}_{cs} = 0.10$). In addition, when we tested the mediation effects of both perceived technology competence and predictions of others' evaluation simultaneously, as shown in Figure 3, both mediating effects remain significant when predicting both investment intentions ($\hat{a}\hat{b} =$ 0.38, 95% CI [0.25, 0.54], $\hat{a}\hat{b}_{cs} =$ 0.13 and $a\hat{}b\hat{} =$ 0.22, 95% CI [0.10, 0.35], $\hat{a}\hat{b}_{cs} =$ 0.07; $p < 0.001$ respectively) and investment amount ($\hat{a}\hat{b} =$ 8,827.00 and $\hat{a}\hat{b}_{cs} =$ 0.10, 95% CI [5,451.00,12,654.00] and $\hat{a}\hat{b} =$ 5,682, $\hat{a}\hat{b}_{cs} =$ 0.06, 95% CI [2,604.00, 9,116.00], $p < 0.001$, respectively).

Insert Figure 3 about here

In a new national context, the U.S, and with a larger and more diverse sample of full-time employees with managerial experiences, Study 2 replicates the results of Study 1 by showing that evaluation criteria consensus boosts investments. More importantly, Study 2 demonstrates that the positive effect of evaluation criteria consensus on investment is mediated by both firm's perceived technological competence and investors' predictions of others' positive evaluations.

Study 3

Studies 3a and 3b aim to replicate the results of Studies 1 and 2 in the real world to establish the external validity of our findings. Using archival datasets, we tested Hypothesis 1 by examining the relationship between evaluation criteria consensus for a new technology and investment in firms using that technology. To the extent that converging results emerge from both the U.S. (Study 3a) and China (Study 3b), we gain increased confidence in their robustness and generalizability.

Setting

Studies 3a and 3b test Hypothesis 1 in the context of AI technologies. AI is a generic name for technologies that rely on computer programs modeled after human brains to draw inferences from big data. These technologies have been used to develop products in various industrial contexts such as security, healthcare, and finance. For example, computer visioning is an AI technology that trains algorithms to recognize and classify different types of objects. The technology has been used to develop intelligent cameras used in surveillance systems.

We chose AI as our research context because AI firms' clients and investors face high uncertainty in evaluating the AI-based programs created by different firms. Partly to resolve uncertainties in evaluation, academic and industry experts have organized technology competitions at academic conferences since the early 2000s. These competitions aimed at facilitating comparisons between different AI-based programs that targeted the same technological problem. For example, an annual competition[iv] that assesses computer programs performance in detecting and classifying objects started in 2010. The evaluation metrics proposed in these competitions influence both academia and industry. For example, in 2015, Microsoft Research cited their performance on a metric (namely top-5 test error) proposed in the above-mentioned competition, to claim that they have surpassed human capacity in classifying images (He et al. 2015).

Data and Sample

Our sample consists of AI firms in the U.S. (Study 3a) and China (Study 3b) because AI firms in these two countries have been shown to receive over 80% of investments to all AI firms worldwide (Insights 2019). To examine the relationship between evaluation criteria consensus and investments, we utilized data from several sources.

First, we collected all evaluation metrics proposed at all 434 technology competitions at 19 top AI-themed academic conferences between 2003, when the first competition was held, and 2019.

The competitions involve four major technological fields of AI: computer visioning and pattern recognition, natural language processing, machine learning, and robotics. To create the list of top AI conferences, we compiled a list of AI conferences that were ranked top 20 in each of the four technology categories according to their h5-index.[v] We collected another list of "top-tier recommended conferences" by China Computer Federation. We obtained the overlapping conferences in these two lists as our final list.

Second, we collected data on investment in AI firms from two major business information platforms, Crunchbase, for firms headquartered in the U.S. (Hallen et al. 2014; Nuscheler et al. 2019), and JuziIT, for firms in China. We chose Crunchbase instead of other venture business information platforms because it contains more comprehensive information on start-ups, particularly the early-stage ones (Gallagher 2013; Hallen 2008). Similarly, we chose JuziIT because it has been widely cited as the most complete and trustworthy information source for AI firms in China (Scmpnews 2017; Technode 2017). We included all firms identified as AI-focused by each platform. We also obtained data on firm characteristics from both platforms. These data were collected by Crunchbase's and JuziIT's data-scraping programs and contributed by firms and investors. They are regularly verified and updated by the platforms. Finally, we obtained data on firm patents from the database of the United States Patents and Trademark Office and the China National Intellectual Property Administration.

Measures

Dependent variable. ***Investment.*** Following previous studies (Pontikes 2012; Wry et al. 2014), we measured investment in a firm with a binary variable, *investment,* with 1 indicating that the firm received investment(s) in a particular year and 0 meaning no investment in that year. We decided not to use investment amount because these figures tend to be inaccurate and often exaggerated (Wry et al. 2014).

Independent variable. ***Evaluation criteria consensus.*** Evaluation criteria consensus is conceptualized as the overall similarity among evaluation criteria proposed in different competitions for evaluating a technology. Following Hsu et al. (2012), we operationalized the construct in four steps. Please See Figure 4 for a demonstration. First, we obtained evaluation metrics proposed in all competitions. Two computer scientists with PhD training calibrated different expressions of the same metric. Second, we compared each competition with all other competitions in the same technology field in each year to calculate the Jaccard similarity score (Bikard 2018; Bikard et al. 2019; Niwattanakul et al. 2013). This indicates the degree of overlap between the two sets of metrics proposed in the two competitions (Online Appendix D provides details about Jaccard similarity score calculation). Third, we summed all pairwise Jaccard similarity scores and divided it by the number of pair-wise comparisons to get the average Jaccard score for the technology. Finally, for each firm, we calculated *evaluation criteria consensus* by summing Jaccard scores for the technology(ies) the firm identifies with and dividing the sum by the number of claimed technology(ies). We matched each firm to one or multiple of the four technology(ies) based on the technology key words that Crunchbase/JuziIT identified. To allow time for the market to respond, we lagged this variable for one year. Table 1shows the keyword matching.

Insert Figure 4 about here

Insert Table 1 about here

Control variables. We controlled for firm-level and field-level variables that may influence investments in a firm. At the firm level, we controlled for firm characteristics that may make investors favor a firm or make a firm self-select into a technology field. First, we controlled for *log(number of patents)* (Hsu and Ziedonis 2008), *number of founders with a PhD*, and *number of founders with science or engineering degree(s)* (Eesley et al. 2014; Hsu 2007) to capture signals about a firm's technological competence. Such signals may affect investment decisions (Busenitz et al. 2005; Conti et al. 2013; Hoenig and Henkel 2015) and the likelihood that a firm will enter different technology fields at different times.

We also controlled for the prestige of founders' universities and past employers, measured by *the number of founders from prestigious universities* (Eesley et al. 2014; MacMillan et al. 1985; Navis and Glynn 2011; Tyebjee and Bruno 1984) and *the number of founders from prestigious companies* (Burton et al. 2002) to capture important status markers found to influence investment decisions and the self-selection of ventures. Appendix E provides a list of prestigious universities and companies. We included a dummy variable, *serial entrepreneur*, with 1 indicating at least one of the founders of the firm is a serial entrepreneur, as on the one hand, investors favor firms founded by serial entrepreneurs (Hsu 2007; Wright et al. 1997); on the other hand, serial entrepreneurs may also be more experienced in identifying technology fields with higher consensus. We included a second dummy variable, *C-round or after,* with 1 indicating the firm was at the C-round or after, as a firm's stage of development is associated with both its likelihood of obtaining investments (Davila et al. 2003; Jeng and Wells 2000) and the development of its field. Finally, to account for industry idiosyncrasies, we controlled for the industrial sector(s) associated with a firm on Crunchbase/ITJuzi with 21/23 dummy indicators in the U.S./Chinese sample.

At the field-level, evaluation criteria consensus is intricately intertwined with efforts in organizing criteria-setting competitions, which could also affect investments in firms. Thus, we added a continuous variable, *number of technology competitions*, measured by the number of competitions involving any of a firm's technologies in a year in our models. To allow time for the market to respond, we lagged this variable for one year.

Estimation

Following past research (Hallen et al. 2014; Wry and Lounsbury 2013), we modeled the effects of evaluation criteria consensus on firms' likelihood of obtaining investments using Cox hazard rate survival models (Cleves et al. 2008; Cox 1972). Our observation window begins in 1999, before which there had been fewer than five investments in any AI firms in U.S./China. Firms founded after 1999 enter the risk-set in their founding years. Firms leave the risk-set in the year they closed or were acquired. We analyzed the relationship between evaluation criteria consensus and investment from 2004 (the year following the first technology competition) to 2019. We lagged our two competition-related variables, evaluation criteria consensus and number of technology competitions for one year. As the data contains repeated observations of firms, we clustered standard errors by firms. To account for the temporal effects, we included year-fixed effects. There are 27,837 and 6,140 firm-yearly observations from 4,905 and 1,345 firms in the U.S. and Chinese sample respectively. We conducted all analyses using Stata 14.2 with the stcox command.

Results

Table 2 presents the descriptive statistics and correlations for the variables included in our analysis. Variance inflation factor (VIF) tests show that none of the models we estimated had VIF scores over 10, the recommended threshold value (Gujarati and Porter 2003).

Insert Table 2 about here

Hypothesis testing. Table 3a presents results from Cox hazard rate regression analysis of investment in AI firms in the U.S. sample. For ease of interpretation and comparison, standardized coefficients are presented. Models 1a and 1b investigate the effects of control variables. Most of the firm-level controls have significant effects on investment in our expected direction in both samples. At the field level, the number of technology competitions has non-significant effects on investment in both samples. This suggests that more efforts at proposing evaluation criteria do not necessarily lead to more investments in firms.

Models 2a and 2b test Hypothesis 1, which predicts evaluation criteria consensus to have a positive effect on investment. In both, evaluation criteria consensus has a significant, positive effect on investment (*Standardized beta* = 0.02, $p < 0.05$ in the U.S. sample and *Standardized beta* = 0.06, $p < 0.001$ in the Chinese sample).

Insert Table 3 about here

Robustness checks. We conducted additional analyses to examine alternative explanations. Table 4 presents our results. First, we ruled out the explanation that the positive effect of evaluation criteria consensus is driven by the growth of a technology field. This logic is in line with population ecologists' claim that growth in an emerging field increases its visibility and legitimacy, making all firms in the field more appealing to investors (Hannan and Carroll 1992; Hannan et al. 1995). Addressing this concern, we controlled for field growth in two different ways. In Model 3a in Table 4a (the U.S. sample) and Model 3b in Table 4b (the Chinese sample), we replaced year dummies with a continuous variable, *year,* to capture the passage of time as a proxy of field growth. Supporting Hypothesis 1, the positive effect of evaluation criteria consensus remained substantial and significant in both models (*Standardized beta* = 0.05, $p < 0.001$ in the U.S. sample and *Standardized beta* = 0.03, $p < 0.001$ in the Chinese sample). In Model 4a and 4b, we controlled for *the number of organizations using a focal firm's technology(ies) (mean-centered)* and the squared term of this variable to capture more nuanced field-growth dynamics (Hannan et al. 1995). Because these two variables are highly correlated with the year dummies and yield VIF scores over 10 (Gujarati and Porter 2003) in the U.S. sample, we removed year dummies in Model 4a. Again, supporting Hypothesis 1, the positive effect of evaluation criteria consensus remained positive and significant in both models (*Standardized beta* = 0.04, $p < 0.001$ in the U.S. sample and *Standardized beta* = 0.02, $p < 0.05$ in the Chinese sample).

Second, we considered the possibility that the observed effect of evaluation criteria consensus on investment is driven by investors' negative reactions to firms associated with a fuzzy category or with multiple categories. We first examined the assumption behind this alternative explanation: firms in fuzzy categories or span categories are those that use technologies with lower evaluation criteria consensus. We constructed measures of *category fuzziness* and *the number of categories a focal firm claims* following Pontikes (2012). Please see Appendix F for details. We control for category fuzziness in Models 5a and 5b (*Standardized beta* = 0.02, $p < 0.05$ in the U.S. sample and *Standardized beta* = 0.06, $p < 0.001$ in the Chinese sample), and the number of categories a focal firm claims in Models 6a and 6b (*Standardized beta* = 0.02, $p < 0.05$ in the U.S. sample and *Standardized beta* = 0.02, $p < 0.05$ in the Chinese sample). The positive effects of evaluation criteria consensus remained significant across the four models, reconfirming Hypothesis 1. Meanwhile, the results also lend modest support to Pontikes's (2012) findings that investors in emerging technology markets favor firms that span multiple categories (Model 5b: *Standardized beta* = 0.02, $p < 0.05$).

Insert Table 4 about here

One potential problem with the Cox hazard rate survival model is that it assumes that the base rate of receiving investment to be constant overtime. The piecewise exponential hazard rate model allows the base rate of receiving funding to vary and thus does not hold a strong assumption about the form of time dependence. Hence, in our third set of robustness checks, we replicated our analysis with piecewise exponential hazard rate models. Results are shown in Table 5. Models 8a and 8b show positive and significant effects of evaluation criteria consensus on investments (*Standardized beta* = 0.29, $p < 0.01$ in the U.S. sample and *Standardized beta* = 0.24, $p < 0.05$ in the Chinese sample).

Insert Table 5 about here

Additional analyses. We also explored how the effect of evaluation criteria consensus might vary across time and for different investors and investment syndicates. Our archival datasets allow us to examine whether the effect of evaluation criteria consensus changes over time. Two processes may induce changes. One is the co-evolution of evaluation criteria consensus and investment. Although investors did not directly participate in the development of technology evaluation criteria (academic and industry scientists did), over time, increasing investments in a technology field may attract standard-setting efforts from technology experts with diverse backgrounds. These diverse experts may propose competing evaluation criterion (Grodal 2018), thus reducing the correlation between investment and evaluation criteria consensus. A second process involves investor learning. As time goes by, investors may accumulate experience and domain-specific knowledge in emerging technology fields, reducing uncertainty about how to assess firms in the sector, and relying less on field-level signals such as evaluation criteria consensus. In sum, both processes imply that the positive effect of evaluation criteria consensus would decline overtime.

We present results of analyses examining these potential effects in Tables 6a (the U.S. sample) and 6b (the Chinese sample). In both models, we used the continuous variable, *year,* to capture the passage of time. The negative, significant coefficients of the interaction term for evaluation consensus and time in Models 9a and 9b indicate that the impact of expert evaluation consensus on investors does decline with time (*Standardized beta* = -12.55, $p < 0.05$ in the U.S. sample and *Standardized beta* = -18.46, $p < 0.001$ in the Chinese sample). Hence, the positive impact of evaluation criteria consensus is the strongest when a technology field is newly formed.

Insert Table 6 about here

The effect of evaluation criteria consensus may also vary for different types of investors. Experienced investors are likely to have routines that help them quickly develop domain-specific knowledge about an emerging field (Cohen 1991; Cohen and Levinthal 1990; Levitt and March

1988). Similarly, professional venture capital firms are apt to have better routines for assessing the value of start-up firms than other types of investors (e.g., public institutions or individual investors) (Gompers and Lerner 2000; Sykes 1990). Therefore, the positive effect of evaluation criteria consensus should be weaker for more experienced investors and professional venture capital firms. We tested these predictions with the U.S. sample and present the results in Table 7. *Investor experience* is a continuous variable measuring the average years of experience among the investors that invest in a firm in a year. Years of experience is calculated by the difference between the year of observation and the founding year of the funding organization. *Professional VC firm* is a dummy variable with 1 indicating that there are one or more professional venture capital firms among the investors that invest in a firm in a year.

In Model 10, the significant, negative coefficient of the interaction term for evaluation criteria consensus and investor experience suggests that the effect of evaluation criteria consensus weakens as the average investment experience of investors increases (*Standardized beta* = -0.01, $p < 0.01$). Similarly, in Model 11, the significant, negative coefficient of the interaction term indicates that the investments by professional VCs are less influenced by expert evaluation consensus (Standardized beta = -0.02, $p < 0.05$).

Third, we explore how the effect of evaluation criteria varies investment syndicates with different sizes. We theorized that investors' concerns over others' opinions was one of the reasons why they rely on field-level evaluation criteria consensus to make investments. In large investment syndicates, because more parties need to be persuaded, the concerns over others' opinions would be stronger. Therefore, we predict the effect of evaluation criteria consensus to be stronger for larger syndicates. We tested this prediction and presented the results in Table 7 (Model 12). *Syndicate size* is a continuous variable measuring the average syndicate size of the investments that a firm received in a year. The positive and significant coefficient of the interaction term of evaluation criteria consensus and syndicate size suggests that the effect of evaluation criteria consensus is stronger for larger syndicates, confirming our prediction (*Standardized beta* = 0.02, $p < 0.001$).

Discussion

This paper investigated an understudied phenomenon in emerging technology fields: how consensus among experts over the criteria to use in evaluating new technologies affect investors' willingness to support firms in the field. We theorized that consensus on evaluation criteria increases willingness to invest by enhancing investors' perceptions of a firm's technological competence and confidence that other investors will view firms in the field positively. We found support for these predictions using data from with two experiments and two archival studies both in the U.S. and in China. The consistency of findings across different methodologies and national contexts boost the validity and generalizability of our effects.

Theoretical Contributions

Our findings have implications for three on-going streams of organizational research on the emergence of new market sectors: the impact of field-level consensus, micro-foundations of the institutionalization processes, and technological evolution.

The impact of field-level consensus. A substantial amount of research has shown that institutions (Grodal 2007, 2018; Hiatt et al. 2009; Lee et al. 2010; Lee et al. 2018; Sine and David 2010) and individual/collective actions (Hannah and Eisenhardt 2018; McDonald and Eisenhardt 2019; Ozcan and Gurses 2018) can reduce the inherent uncertainties of an emerging sector (Dattée et al. 2018; Santos and Eisenhardt 2009). Studies have shown that field-level consensus on industry architecture (Ozcan and Santos 2015), producer identities (Georgallis et al. 2018), and field/technology/industry frames (Gurses and Ozcan 2015; Hiatt and Carlos 2019; Kaplan 2008; Lounsbury et al. 2003) contribute to the formation and stabilization of an emerging field. Yet this line of research has not examined the impact of field-level consensus in evaluation criteria among technology experts, an important cognitive institution in a new market sector. This study not only presents the first concrete measure of expert evaluation criteria consensus over time and across different technology fields, but also theorized and tested its impact on investment to an emerging sector. By doing so, we advance existing research on evaluation institutions in emerging market sectors (Goldfarb et al. 2018; Rao 1994; Sine et al. 2007), which has not explicitly considered potential disagreements in evaluation criteria (Grodal 2018; Lee et al. 2017). We show that experts' consensus on technology evaluation criteria induces positive evaluations of and draws investments to ventures in an emerging market sector.

Micro-foundations of the institutionalization processes. While existing research has examined how the development of cognitive institutions in an emerging field (Grodal et al. 2015; Kennedy 2008; Lounsbury and Glynn 2001; Navis and Glynn 2010; Wry et al. 2011) affects individual firms and the entire field (Hiatt et al. 2009; Lounsbury and Glynn 2001; Sine and Lee 2009; Wry et al. 2011), few studies have captured the underlying micro-processes. A crucial micro-process in this vein is how investors evaluate and support new organizations. Building on works on investor sense-making (Navis and Glynn 2011; Weber and Glynn 2006), we develop and test a theory positing that consensus among experts in evaluation criteria—a crucial field-level cognitive institution (Khaire and Wadhwani 2010)—boosts investors' assessment of firms in emerging technology fields. We found two mediating mechanisms, namely perceived technological competence and prediction of other's positive evaluations. Finding perceived technological competence as a mediator enriches our knowledge of providers' uncertainty about quality (Navis and Glynn 2011), while showing prediction of others' evaluations as a mediator informs research on uncertainty about other stakeholders (Correll et al. 2017). This paper is one of the first to explicate the micro-level processes through which field-level cognitive institutions (e.g. evaluation criteria consensus) legitimate individual organizations in an emerging sector.

Technological evolution. The literature on technological evolution and industry emergence has a long-standing concern with technology standards (Tassey 2000; Vakili 2016) and dominant designs (Anderson and Tushman 1990; Suárez and Utterback 1995). Researchers show that they reduce uncertainty in a nascent sector by coordinating transactions among market actors. Recently, organizational scholars have begun to examine how the development of cognitive institutions, such as categories (Grodal 2007; Suarez et al. 2015; Zunino et al. 2019), technology frames (Kaplan 2008; Kaplan and Tripsas 2008), and field schemas and labels (Bingham and Kahl 2013), affect uncertainty in an emerging technology field. We join this burgeoning literature to investigate how a relatively understudied cognitive institution—evaluation criteria—influences firm resource acquisition before the existence of agreed-upon technology standards and dominant designs. Our findings indicate that experts reaching consensus in the technology's evaluation criteria constitutes a critical milestone in the development of a field.

Limitations and Future Research Directions

This study has a few limitations which promise future research opportunities. First, this paper focused on the consequences of consensus in evaluation criteria rather than its antecedents. Thus, we chose not to focus on how convergence or divergence in evaluation criteria occur. This question, however, promises future research opportunities on the still mysterious process of development and evolution of cognitive institutions in emerging sectors (Grodal 2018; Hoffman 1999). The co-existing evaluation criterion in an emerging technology field usually represent different underlying technology frames (Kaplan 2008; Kaplan and Tripsas 2008) and are associated with individual/organizational actors with different vested interests (Campbell 1998; Hoffman 1999). For example, based on our data, the AI field witnessed a co-evolution of the evaluation criteria and the organizations that proposed these criteria in competitions. From the mid-2000s until 2013, many early competitions were organized by a core group of academics and industry researchers who inherited past evaluation criteria. In 2013, the number of competitions doubled. After 2013, some competitions were entirely organized by technology companies, many of which proposed new criteria. Such changes fueled disagreements over the proper criteria to use in evaluating technology solutions. We call for future investigation of the co-evolution of organizer networks and evaluation criteria. Studying this would not only answer the intriguing question of how consensus arises, but also unveil the co-evolution of networks, power dynamics, and institutions (Padgett and Powell 2012; Powell and Oberg 2017).

Second, we focus on the effect of field-level consensus in technology evaluation criteria on investments in all firms in the field. We chose this focus because this paper, as the first theorization and test of evaluation criteria consensus, aims to document its effect on the entire emerging sector. Nevertheless, we tested factors that may create within-field variance in our effects. For example, in our archival studies (Studies 3a and 3b), we explored how the effect varies over time and across investors and syndicates. In our experimental Study A1, we tested how information about a firm's technological competence moderates the positive effects of evaluation criteria consensus (See Appendix G for details). These additional analysis points to how characteristics of firms, technologies, industries, and investors cause heterogeneity in the effects of evaluation criteria consensus. We encourage future researchers to test other potential moderators. For example, substantial differences in resources, competences, and legitimacy between de alio and de novo firms (Carroll et al. 1996; Helfat and Lieberman 2002; Khessina and Carroll 2008), may influence the extent to which they benefit from field-level consensus in evaluation criteria. Due to our focus on start-up firms, we did not examine this question. On a separate note, consensus in evaluation criteria resolves two types of uncertainties in an emerging technology field—technical uncertainty (i.e. what features to measure) and standard uncertainty (i.e. what threshold to use) (Graffin and Ward 2010). The two types of uncertainties do not always co-vary. In other words, one field may have low technical uncertainty (the technology performance is perfectly measurable) but high standard uncertainty (the threshold for good performance is unclear), whereas another field may have the opposite. We encourage future research on how these two types of uncertainties and their combinations moderate the effect of evaluation criteria consensus.

Conclusion

Processes of industry/market/field/category emergence have attracted considerable scholarly effort in the past decade (e.g. Grodal et al. 2015; Kennedy 2008; Padgett and Powell 2012). While a substantial portion of these works examined the social evaluation processes in nascent fields and their impact on organizational survival and performance (e.g. Pontikes 2012; Wry et al. 2014), little attention has been paid to the field-level institutions that facilitate or prohibit evaluation. Our study, which focused on the consequence of consensus on evaluation criteria on firms' ability to obtain investments, is a step towards understanding developments of evaluation institutions. We hope to stimulate future research that digs deeper into processes and consequences of evaluation institutions in emerging fields.

References

Aggarwal R, Kryscynski D, Singh H (2015) Evaluating venture technical competence in venture capitalist investment decisions. *Management Sci.* 61(11): 2685-2706.

Aldrich HE, Fiol CM (1994) Fools rush in? The institutional context of industry creation. *Acad. Management Rev* 19(4): 645-670.

Anderson P, Tushman ML (1990) Technological discontinuities and dominant designs: A cyclical model of technological change. *Admin. Sci. Quar.* 35(4): 604-633.

Baum JA, Silverman BS (2004) Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *J. Bus. Venturing* 19(3): 411-436.

Biglaiser G (1993) Middlemen as experts. *The RAND journal of Economics* 212-223.

Biglaiser G, Friedman JW (1994) Middlemen as guarantors of quality. *International journal of industrial organization* 12(4): 509-531.

Bikard M (2018) Made in academia: The effect of institutional origin on inventors' attention to science. *Organ. Sci.* 29(5): 818-836.

Bikard M, Vakili K, Teodoridis F (2019) When collaboration bridges institutions: The impact of university–industry collaboration on academic productivity. *Organ. Sci.* 30(2): 426-445.

Bingham CB, Kahl SJ (2013) The process of schema emergence: Assimilation, deconstruction, unitization and the plurality of analogies. *Acad. Management J.* 56(1): 14-34.

Bitektine A, Lucas JW, Schilke O (2018) Institutions under a microscope: Experimental methods in institutional theory. *Unconventional Methodology in Organization and Management Research* (Oxford University Press), 147-167.

Burton MD, Sørensen JB, Beckman CM (2002) Coming from good stock: Career histories and new venture formation. *Social structure and organizations revisited* (Emerald Group Publishing Limited), 229-262.

Busenitz LW, Fiet JO, Moesel DD (2005) Signaling in Venture Capitalist—New Venture Team Funding Decisions: Does it Indicate Long–Term Venture Outcomes? *Entrepreneurship Theory and Practice* 29(1): 1-12.

Campbell JL (1998) Institutional analysis and the role of ideas in political economy. *Theory and society* 27(3): 377-409.

Carroll GR, Bigelow LS, Seidel MDL, Tsai LB (1996) The fates of de novo and de alio producers in the American automobile industry 1885–1981. *Strategic Management J* 17(S1): 117-137.

Chatterji AK, Durand R, Levine DI, Touboul S (2016) Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management J* 37(8): 1597-1614.

Cleves M, Gould W, Gould WW, Gutierrez R, Marchenko Y (2008) *An introduction to survival analysis using Stata* (Stata press).

Cohen MD (1991) Individual learning and organizational routine: Emerging connections. *Organ. Sci.* 2(1): 135-139.

Cohen WM, Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Quar.* 35(1): 128-152.

Conti A, Thursby M, Rothaermel FT (2013) Show me the right stuff: Signals for high ‐ tech startups. *Journal of Economics & Management Strategy* 22(2): 341-364.

Correll SJ, Ridgeway CL, Zuckerman EW, Jank S, Jordan-Bloch S, Nakagawa S (2017) It's the conventional thought that counts: How third-order inference produces status advantage. *Amer. Sociol. Rev.* 82(2): 297-327.

Cox DR (1972) Regression models and life ‐ tables. *Journal of the Royal Statistical Society: Series B*

(Methodological) 34(2): 187-202.
Cumming D, binti Johan SA (2008) Preplanned exit strategies in venture capital. *Eur. Econ. Rev.* 52(7): 1209-1241.
Dattée B, Alexy O, Autio E (2018) Maneuvering in poor visibility: How firms play the ecosystem game when uncertainty is high. *Acad. Management J.* 61(2): 466-498.
Davila A, Foster G, Gupta M (2003) Venture capital financing and the growth of startup firms. *J. Bus. Venturing* 18(6): 689-708.
Dew N, Read S, Sarasvathy SD, Wiltbank R (2009) Effectual versus predictive logics in entrepreneurial decision-making: Differences between experts and novices. *J. Bus. Venturing* 24(4): 287-309.
Durand R, Kremp P-A (2016) Classical deviation: Organizational and individual status as antecedents of conformity. *Acad. Management J.* 59(1): 65-89.
Durand R, Paolella L (2013) Category stretching: Reorienting research on categories in strategy, entrepreneurship, and organization theory. *J. Manag. Stud.* 50(6): 1100-1123.
Eesley CE, Hsu DH, Roberts EB (2014) The contingent effects of top management teams on venture performance: Aligning founding team composition with innovation strategy and commercialization environment. *Strategic Management J* 35(12): 1798-1817.
Fisher G, Kuratko DF, Bloodgood JM, Hornsby JS (2017) Legitimate to whom? The challenge of audience diversity and new venture legitimacy. *J. Bus. Venturing* 32(1): 52-71.
Florin J, Lubatkin M, Schulze W (2003) A social capital model of high-growth ventures. *Acad. Management J.* 46(3): 374-384.
Gallagher B (2013) How CrunchBase Data Compares To Other Industry Sources. *Tech Crunch* (July 23), http://social.techcrunch.com/2013/07/23/how-crunchbase-data-compares-to-other-industry-sources/.
Georgallis P, Dowell G, Durand R (2018) Shine on me: Industry coherence and policy support for emerging industries. *Admin. Sci. Quar.* 64(3): 503-541.
Goldfarb B, Zavyalova A, Pillai S (2018) Did victories in certification contests affect the survival of organizations in the American automobile industry during 1895–1912? A replication study. *Strategic Management J* 39(8): 2335-2361.
Gompers P, Lerner J (2000) The determinants of corporate venture capital success: Organizational structure, incentives, and complementarities. *Concentrated corporate ownership* (University of Chicago Press), 17-54.
Graffin SD, Ward AJ (2010) Certifications and reputation: Determining the standard of desirability amidst uncertainty. *Organ. Sci.* 21(2): 331-346.
Grodal S (2007) The emergence of a new organizational field: Labels, meaning and emotions in nanotechnology. Unpublished doctoral dissertation, Stanford University Palo Alto, California.
Grodal S (2018) Field expansion and contraction: How communities shape social and symbolic boundaries. *Admin. Sci. Quar.* 63(4): 783-818.
Grodal S, Gotsopoulos A, Suarez FF (2015) The coevolution of technologies and categories during industry emergence. *Acad. Management Rev* 40(3): 423-445.
Gujarati DN, Porter DC (2003) *Basic Econometrics*, 4th ed. (McGraw-Hill, New York).
Gurses K, Ozcan P (2015) Entrepreneurship in regulated markets: framing contests and collective action to introduce pay TV in the US. *Acad. Management J.* 58(6): 1709-1739.
Hallen BL (2008) The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? *Admin. Sci. Quar.* 53(4): 685-718.
Hallen BL, Bingham CB, Cohen S (2014) Do accelerators accelerate? A study of venture accelerators as a path to success? *Academy of Management Proceedings* (Academy of Management Briarcliff Manor, NY 10510), 12955.
Hannah DP, Eisenhardt KM (2018) How firms navigate cooperation and competition in nascent ecosystems.

Strategic Management J 39(12): 3163-3192.
Hannan MT, Carroll GR (1992) *Dynamics of organizational populations: Density, legitimation, and competition* (Oxford University Press).
Hannan MT, Carroll GR, Dundon EA, Torres JC (1995) Organizational evolution in a multinational context: Entries of automobile manufacturers in Belgium, Britain, France, Germany, and Italy. *Amer. Sociol. Rev.* 60(4): 509-528.
Hayes AF (2017) *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Guilford Publications).
He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imagenet classification *Proceedings of the IEEE international conference on computer vision* (ICCV), 1026-1034.
Helfat CE, Lieberman MB (2002) The birth of capabilities: market entry and the importance of pre‐history. *Ind. Corp. Change* 11(4): 725-760.
Hiatt SR, Carlos WC (2019) From farms to fuel tanks: Stakeholder framing contests and entrepreneurship in the emergent US biodiesel market. *Strategic Management J* 40(6): 865-893.
Hiatt SR, Sine WD, Tolbert PS (2009) From Pabst to Pepsi: The deinstitutionalization of social practices and the creation of entrepreneurial opportunities. *Admin. Sci. Quar.* 54(4): 635-667.
Hoenig D, Henkel J (2015) Quality signals? The role of patents, alliances, and team experience in venture capital financing. *Res. Policy* 44(5): 1049-1064.
Hoffman AJ (1999) Institutional evolution and change: Environmentalism and the US chemical industry. *Acad. Management J.* 42(4): 351-371.
Hsu DH (2007) Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Res. Policy* 36(5): 722-741.
Hsu DH, Ziedonis RH (2008) Patents as quality signals for entrepreneurial ventures *Academy of Management Proceedings* (Academy of Management Briarcliff Manor, NY 10510), 1-6.
Hsu G, Roberts PW, Swaminathan A (2012) Evaluative schemas and the mediating role of critics. *Organization Science* 23(1): 83-97.
Insights C (2019) China Is Starting To Edge Out The US In AI Investment. *CB Insights Research* (December 2), https://www.cbinsights.com/research/china-artificial-intelligence-investment-startups-tech/.
Jeng LA, Wells PC (2000) The determinants of venture capital funding: evidence across countries. *J. Corp. Finan.* 6(3): 241-289.
Jensen M (2006) Should we stay or should we go? Accountability, status anxiety, and client defections. *Admin. Sci. Quar.* 51(1): 97-128.
Kanze D, Huang L, Conley MA, Higgins ET (2018) We ask men to win and women not to lose: Closing the gender gap in startup funding. *Acad. Management J.* 61(2): 586-614.
Kaplan S (2008) Framing contests: Strategy making under uncertainty. *Organ. Sci.* 19(5): 729-752.
Kaplan S, Tripsas M (2008) Thinking about technology: Applying a cognitive lens to technical change. *Res. Policy* 37(5): 790-805.
Kennedy MT (2008) Getting counted: Markets, media, and reality. *Amer. Sociol. Rev.* 73(2): 270-295.
Khaire M, Wadhwani RD (2010) Changing landscapes: The construction of meaning and value in a new market category—Modern Indian art. *Acad. Management J.* 53(6): 1281-1304.
Khessina OM, Carroll GR (2008) Product demography of de novo and de alio firms in the optical disk drive industry, 1983–1999. *Organ. Sci.* 19(1): 25-38.
Kynn M (2008) The 'heuristics and biases' bias in expert elicitation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(1): 239-264.
Lanahan L, Armanios D (2018) Does More Certification Always Benefit a Venture? *Organ. Sci.* 29(5): 931-947.

Lee B, Sine W, Tolbert P (2010) Certifying the harvest: the role of standards-based certification organizations in the organic food industry. Working paper, Cornell University, Ithaca, NY.

Lee BH, Hiatt SR, Lounsbury M (2017) Market mediators and the trade-offs of legitimacy-seeking behaviors in a nascent category. *Organ. Sci.* 28(3): 447-470.

Lee BH, Struben J, Bingham CB (2018) Collective action and market formation: An integrative framework. *Strategic Management J* 39(1): 242-266.

Levitt B, March JG (1988) Organizational learning. *Annu. Rev. Sociol.* 14(1): 319-338.

Lounsbury M, Glynn MA (2001) Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management J* 22(6 - 7): 545-564.

Lounsbury M, Ventresca M, Hirsch PM (2003) Social movements, field frames and industry emergence: a cultural–political perspective on US recycling. *Socio-economic review* 1(1): 71-104.

MacMillan IC, Siegel R, Narasimha PS (1985) Criteria used by venture capitalists to evaluate new venture proposals. *J. Bus. Venturing* 1(1): 119-128.

MacMillan IC, Zemann L, Subbanarasimha P (1987) Criteria distinguishing successful from unsuccessful ventures in the venture screening process. *J. Bus. Venturing* 2(2): 123-137.

Mason C, Stark M (2004) What do investors look for in a business plan? A comparison of the investment criteria of bankers, venture capitalists and business angels. *Int. Small Bus. J.* 22(3): 227-248.

McDonald RM, Eisenhardt KM (2019) Parallel play: Startups, nascent markets, and effective business-model design. *Admin. Sci. Quar.* 0001839219852349.

McKendrick DG, Jaffee J, Carroll GR, Khessina OM (2003) In the bud? Disk array producers as a (possibly) emergent organizational form. *Admin. Sci. Quar.* 48(1): 60-93.

Moeen M, Agarwal R (2017) Incubation of an industry: Heterogeneous knowledge bases and modes of value capture. *Strategic Management J* 38(3): 566-587.

Morgan MG (2014) Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National academy of Sciences* 111(20): 7176-7184.

Navis C, Glynn MA (2010) How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. *Admin. Sci. Quar.* 55(3): 439-471.

Navis C, Glynn MA (2011) Legitimate distinctiveness and the entrepreneurial identity: Influence on investor judgments of new venture plausibility. *Acad. Management Rev* 36(3): 479-499.

Negro G, Hannan MT, Rao H (2010) Categorical contrast and audience appeal: Niche width and critical success in winemaking. *Ind. Corp. Change* 19(5): 1397-1425.

Negro G, Hannan MT, Rao H (2011) Category reinterpretation and defection: Modernism and tradition in Italian winemaking. *Organ. Sci.* 22(6): 1449-1463.

Niwattanakul S, Singthongchai J, Naenudorn E, Wanapu S (2013) Using of Jaccard coefficient for keywords similarity *Proceedings of the international multiconference of engineers and computer scientists* (IMECS, Hong Kong), 380-384.

Nuscheler D, Engelen A, Zahra SA (2019) The role of top management teams in transforming technology-based new ventures' product introductions into growth. *J. Bus. Venturing* 34(1): 122-140.

Ozcan P, Gurses K (2018) Playing cat and mouse: Contests over regulatory categorization of dietary supplements in the United States. *Acad. Management J.* 61(5): 1789-1820.

Ozcan P, Santos FM (2015) The market that never was: Turf wars and failed alliances in mobile payments. *Strategic Management J* 36(10): 1486-1512.

Padgett JF, Powell WW (2012) *The emergence of organizations and markets* (Princeton University Press).

Plummer LA, Allison TH, Connelly BL (2016) Better together? Signaling interactions in new venture pursuit of

initial external capital. *Acad. Management J.* 59(5): 1585-1604.
Pollock TG, Gulati R (2007) Standing out from the crowd: The visibility-enhancing effects of IPO-related signals on alliance formation by entrepreneurial firms. *Strategic Organization* 5(4): 339-372.
Pontikes EG (2012) Two sides of the same coin: How ambiguous classification affects multiple audiences' evaluations. *Administrative Science Quarterly* 57(1): 81-118.
Powell WW, Oberg A (2017) Networks and institutions. *The Sage handbook of organizational institutionalism* 446-476.
Preacher KJ, Hayes AF (2004) SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers* 36(4): 717-731.
Rao H (1994) The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912. *Strategic Management J* 15(S1): 29-44.
Santos FM, Eisenhardt KM (2009) Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Acad. Management J.* 52(4): 643-671.
Scmpnews (2017) China turns to artificial intelligence to boost its education system. *South China Morning Post* (October 14), https://www.scmp.com/tech/science-research/article/2115271/china-wants-bring-artificial-intelligence-its-classrooms-boost.
Sine WD, David RJ (2010) Institutions and entrepreneurship. *Institutions and Entrepreneurship* (Emerald Group Publishing Limited), 1-26.
Sine WD, David RJ, Mitsuhashi H (2007) From plan to plant: Effects of certification on operational start-up in the emergent independent power sector. *Organ. Sci.* 18(4): 578-594.
Sine WD, Lee BH (2009) Tilting at windmills? The environmental movement and the emergence of the US wind energy sector. *Admin. Sci. Quar.* 54(1): 123-155.
Suarez FF, Grodal S, Gotsopoulos A (2015) Perfect timing? Dominant category, dominant design, and the window of opportunity for firm entry. *Strategic Management J* 36(3): 437-448.
Suárez FF, Utterback JM (1995) Dominant designs and the survival of firms. *Strategic Management J* 16(6): 415-430.
Suchman MC (1995) Managing legitimacy: Strategic and institutional approaches. *Acad. Management Rev* 20(3): 571-610.
Sykes HB (1990) Corporate venture capital: Strategies for success. *J. Bus. Venturing* 5(1): 37-47.
Tassey G (2000) Standardization in technology-based markets. *Res. Policy* 29(4-5): 587-602.
Technode (2017) China vs US: Who is winning the big AI battle? *Technode China* (October 22), https://technode.com/2017/10/22/china-vs-us-ai/.
Tetlock PE, Skitka L, Boettger R (1989) Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *J. Pers. Soc. Psychol.* 57(4): 632.
Tonoyan V, Strohmeyer R, Jennings JE (2019) Gender gaps in perceived start-up ease: Implications of sex-based labor market segregation for entrepreneurship across 22 European countries. *Admin. Sci. Quar.* 0001839219835867.
Tushman ML (1992) Organizational determinants of technological change: toward a sociology of technological evolution. *Res. Organ. Behav.* 14 311-347.
Tushman ML, Anderson P (1986) Technological discontinuities and organizational environments. *Admin. Sci. Quar.* 31(3): 439-465.
Tyebjee TT, Bruno AV (1984) A model of venture capitalist investment activity. *Management Sci.* 30(9): 1051-1066.
Vakili K (2016) Collaborative promotion of technology standards and the impact on innovation, industry structure,

and organizational capabilities: Evidence from modern patent pools. *Organ. Sci.* 27(6): 1504-1524.
Weber K, Glynn MA (2006) Making sense with institutions: Context, thought and action in Karl Weick's theory. *Organ. Stud.* 27(11): 1639-1660.
Wright M, Robbie K, Ennew C (1997) Venture capitalists and serial entrepreneurs. *J. Bus. Venturing* 12(3): 227-249.
Wry T, Lounsbury M (2013) Contextualizing the categorical imperative: Category linkages, technology focus, and resource acquisition in nanotechnology entrepreneurship. *J. Bus. Venturing* 28(1): 117-133.
Wry T, Lounsbury M, Glynn MA (2011) Legitimating nascent collective identities: Coordinating cultural entrepreneurship. *Organ. Sci.* 22(2): 449-463.
Wry T, Lounsbury M, Jennings PD (2014) Hybrid vigor: Securing venture capital by spanning categories in nanotechnology. *Academy of Management Journal* 57(5): 1309-1333.
Yue LQ, Luo J, Ingram P (2013) The failure of private regulation: Elite control and market crises in the Manhattan banking industry. *Admin. Sci. Quar.* 58(1): 37-68.
Zuckerman EW (1999) The categorical imperative: Securities analysts and the illegitimacy discount. *Amer. J. Sociol.* 104(5): 1398-1438.
Zunino D, Suarez FF, Grodal S (2019) Familiarity, Creativity, and the Adoption of Category Labels in Technology Industries. *Organ. Sci.* 30(1): 169-190.

Figure 1. Summary of Theoretical Mechanisms

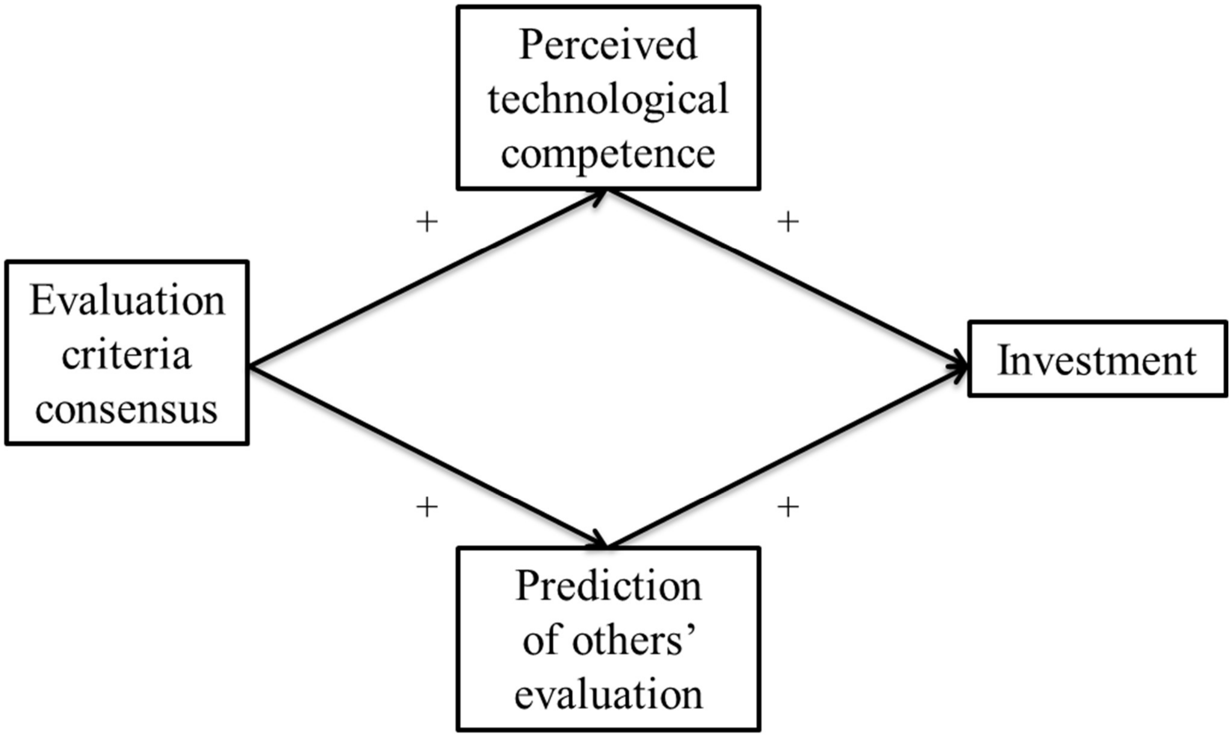
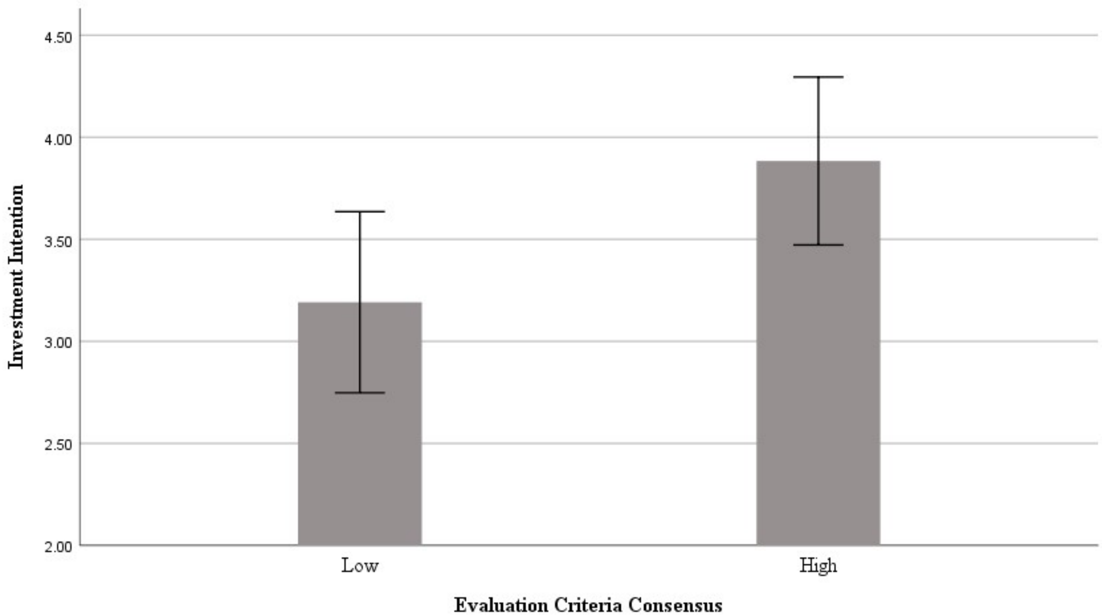
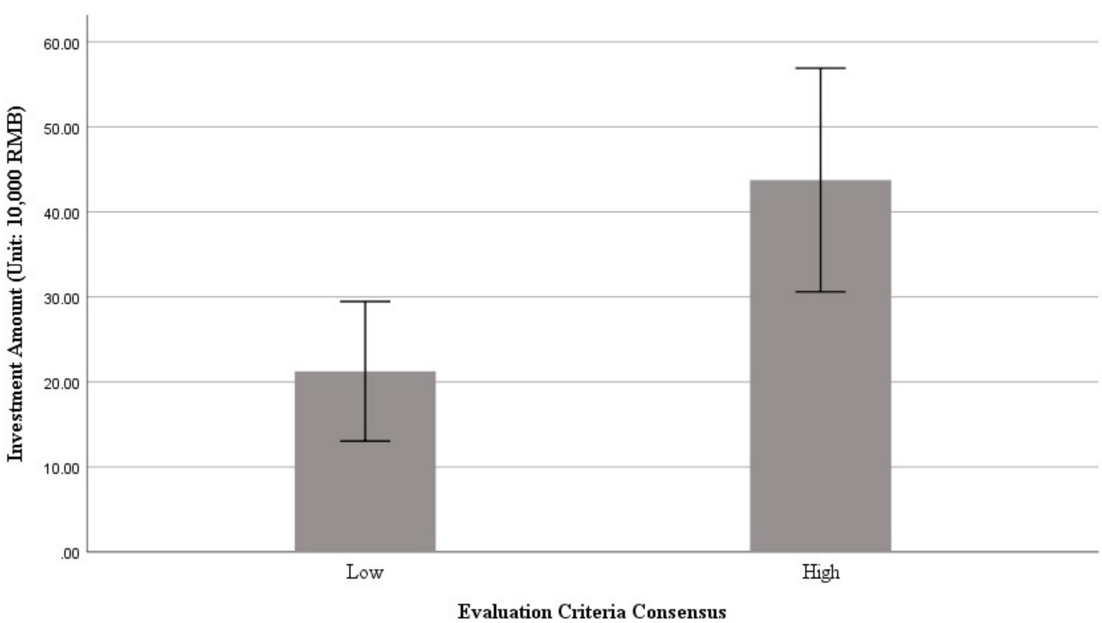


Figure 2. Study 1. Difference in Investment Cross Conditions

(a)



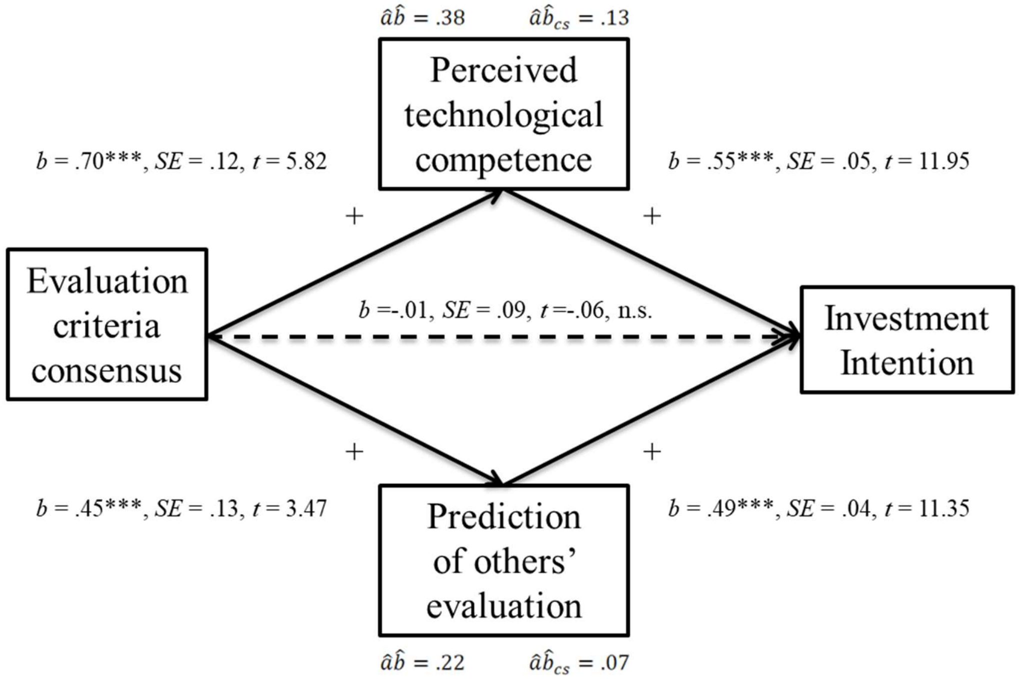
(b)



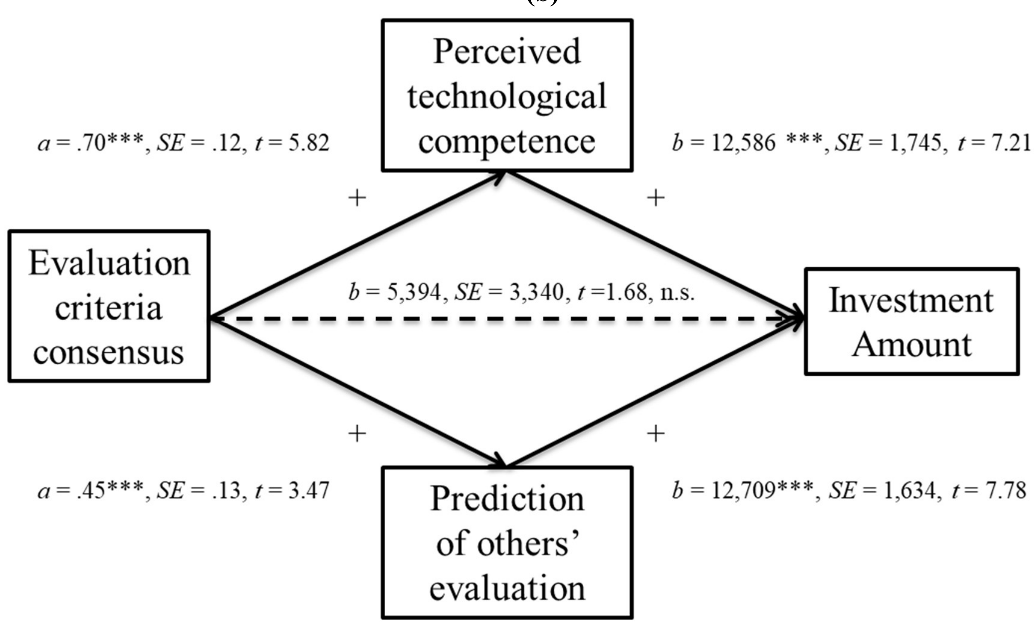
Notes. Panel (a) shows the mean difference cross conditions in investment intentions, and panel (b) shows the mean difference cross conditions in investment amount. Error bars shows 95% CI. N=80.

Figure 3. Study 2. Results of the Mediation Analysis

(a)



(b)



Notes. Panel (a) shows results from mediation analysis on investment intention. Panel (b) shows results from mediation analysis on investment amount. Asterisks indicate significant coefficients. N=412. *** $p < 0.001$.

Figure 4. Study 3. An Example of Calculation of Evaluation Criteria Consensus

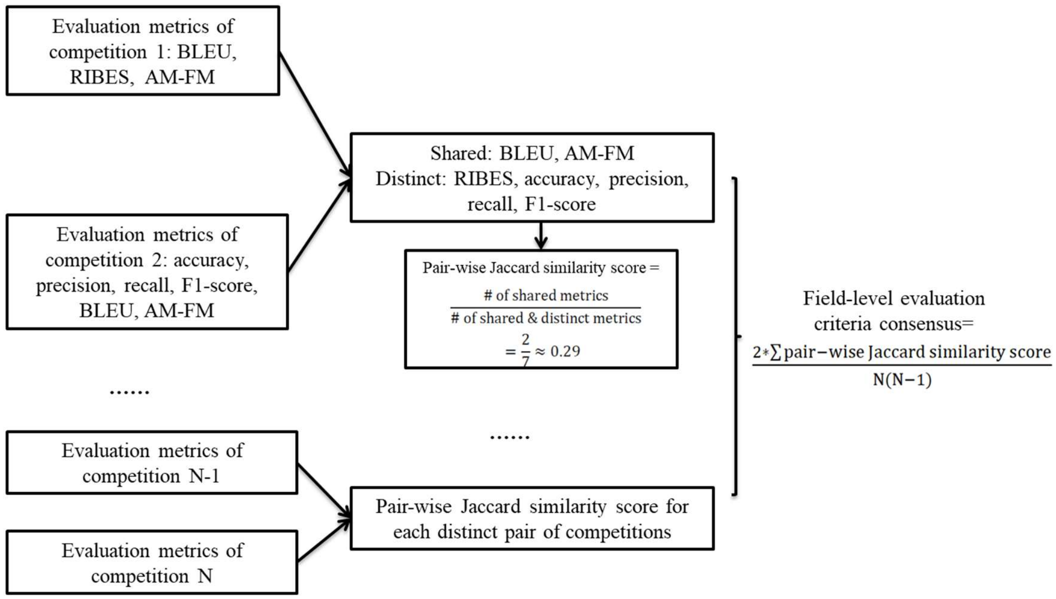


Table 1. Study 3. Matching Firm Technology with Technology Fields

Technology field	Firm technology
Computer vision and pattern recognition	Computer Vision, Image Recognition, Image Processing, Machine Vision, Video, Facial Recognition, Speech Recognition
Machine learning	Machine Learning, Deep Learning, Big Data, Data Mining, Data Analytics, Predictive Analytics
Natural language processing	Natural Language, Translation, Text Analytics
Robotics	Robotics

Table 2. Study 3. Summary Statistics

(a). US Sample (N=27,837)

		Mean	SD	1	2	3	4	5	6	7	8	9
1	*Investment (1=Yes, 0=No)*	0.03	0.17									
2	*Evaluation criteria consensus*	0.33	0.37	0.04								
3	*Number of founders from prestigious universities*	0.18	0.48	0.12	0.04							
4	*Number of founders from prestigious companies*	0.13	0.40	0.09	0.01	0.34						
5	*Number of founders with a PhD*	0.05	0.24	0.08	0.03	0.39	0.17					
6	*Number of founders with science/engineering degree(s)*	0.13	0.41	0.08	0.03	0.40	0.34	0.22				
7	*Serial entrepreneur (1=Yes, 0=No)*	0.06	0.24	0.02	0.01	0.09	0.09	0.05	0.07			
8	*Log(Number of patents)*	0.10	0.53	-0.01	-0.01	0.06	0.08	0.01	0.04	0.00		
9	*C-round or after (1=Yes, 0=No)*	0.00	0.01	0.07	-0.00	0.01	0.00	0.02	0.00	0.00	0.00	
10	*Number of technology competitions*	6.97	9.61	0.00	0.21	0.01	0.03	0.04	0.00	0.02	-0.01	-0.01

(b). Chinese Sample (N=6,140)

		Mean	SD	1	2	3	4	5	6	7	8	9
1	*Investment (1=Yes, 0=No)*	0.14	0.34									
2	*Evaluation criteria consensus*	0.23	0.37	0.15								
3	*Number of founders from prestigious universities*	0.11	0.62	0.20	0.03							
4	*Number of founders from prestigious companies*	0.08	0.50	0.24	0.05	0.19						
5	*Number of founders with a PhD*	0.06	0.30	0.26	0.07	0.48	0.29					
6	*Number of founders with science/engineering degree(s)*	0.02	0.15	0.15	0.06	0.20	0.25	0.29				
7	*Serial entrepreneur (1=Yes, 0=No)*	0.01	0.12	0.11	0.03	0.04	0.16	0.05	0.02			
8	*Log(Number of patents)*	1.61	7.18	0.02	0.07	0.00	- 0.01	0.01	- 0.01	- 0.01		
9	*C-round or after (1=Yes, 0=No)*	0.03	0.16	0.06	- 0.01	0.10	0.08	0.06	- 0.02	0.01	0.02	
10	*Number of technology competitions*	5.40	9.49	0.03	0.21	- 0.02	0.00	0.02	0.03	- 0.01	0.00	- 0.03

Note. SD, Standard deviation. For brevity, industry dummy indicators were omitted. All coefficients with absolute value above 0.05 are significant at $p <$ 0.001. Two-tailed test.

Table 3. Study 3. Cox Hazard Rate Models of Investments to AI Firms

(a). The U.S. sample (N=27,837)

Variable	Model 1a	Model 2a
Evaluation criteria consensus		0.02*
		(0.11)
Firm-level controls		
Number of founders from prestigious universities	0.05***	0.04***
	(0.06)	(0.06)
Number of founders from prestigious companies	0.04***	0.04***
	(0.07)	(0.07)
Number of founders with a PhD	0.01*	0.01*
	(0.09)	(0.09)
Number of founders with science/engineering degree(s)	0.02**	0.02**
	(0.07)	(0.07)
Serial entrepreneur (1=Yes, 0=No)	0.01	0.01
	(0.13)	(0.13)
Log(Number of patents)	-0.01	-0.00
	(0.07)	(0.07)
C-round or after (1=Yes, 0=No)	0.01***	0.01***
	(0.3)	(0.3)
Field-level controls		
Number of technology competitions	0.01	0.00
	(0.00)	(0.00)
Industry dummies	Yes	Yes
Year fixed effects	Yes	Yes
Log pseudolikelihood	-6577.13	-6674.58

(b). The Chinese sample (N=6,140)

Variable	Model 1b	Model 2b
Evaluation criteria consensus		0.06***
		(0.09)
Firm-level controls		
Number of founders from prestigious universities	0.02***	0.02***
	(0.02)	(0.02)
Number of founders from prestigious companies	0.02***	0.02***
	(0.03)	(0.03)
Number of founders with a PhD	0.03***	0.03***
	(0.05)	(0.05)
Number of founders with science/engineering degree(s)	0.02	0.01
	(0.14)	(0.13)
Serial entrepreneur (1=Yes, 0=No)	0.02***	0.02***
	(0.12)	(0.11)
Log(Number of patents)	0.02***	0.02**
	(0.00)	(0.00)
C-round or after (1=Yes, 0=No)	0.02**	0.02***
	(0.13)	(0.132
Field-level controls		
Number of technology competitions	0.01	0.01
	(0.00)	(0.00)
Industry dummies	Yes	Yes
Year fixed effects	Yes	Yes
Log pseudolikelihood	-5407.95	-5381.45

Note. Industry dummies are omitted for brevity. Standardized beta coefficients are reported. Standard error in parentheses. All two-tailed tests. Robust standard error clustered by firm. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 4. Study 3. Robustness Checks: Cox Hazard Rate Models of Investments to AI Firms

(a). The U.S. sample (N=27,837)

Variable	Model 3a	Model 4a	Model 5a	Model 6a
Evaluation criteria consensus	0.05***	0.04***	0.02*	0.02*
	(0.10)	(0.12)	(0.11)	(0.11)
Year	0.01			
	(0.01)			
Number of organizations using a focal firm's technology(ies) (Mean-centered)		0.09***		
		(0.00)		
Number of organizations using a focal firm's technology(ies) (Mean-centered, squared)		-0.09***		
		(0.00)		
Number of categories focal firm claims			0.02	
			(0.02)	
Category fuzziness				0.00
				(2.46)
Controls	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes
Log pseudolikelihood	-6618.76	-6609.54	-6574.58	-6576.11

(b). The Chinese sample (N=6,140)

Variable	Model 3b	Model 4b	Model 5b	Model 6b
Evaluation criteria consensus	0.03***	0.02*	0.06***	0.02*
	(0.09)	(0.12)	(0.08)	(0.10)
Year	0.13***			
	(0.01)			
Number of organizations using a focal firm's technology(ies) (Mean-centered)		0.12***		
		(0.00)		
Number of organizations using a focal firm's technology(ies) (Mean-centered, squared)		-0.04**		
		(0.00)		
Number of categories focal firm claims			0.02**	
			(0.24)	
Category fuzziness				0.03
				(12.31)
Controls	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes
Log pseudolikelihood	-5315.09	-5365.87	-5195.02	-5378.79

Note. Control variables and industry dummies are omitted for brevity. Standardized beta coefficients are reported. Standard errors are in parentheses. All two-tailed tests. Robust standard error clustered by firm. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 5. Study 3. Robustness Checks: Piecewise Hazard Rate Models of Investments to AI Firms

(a). The U.S. sample (N=27,837)

Variable	Model 7a	Model 8a
Evaluation criteria consensus		0.29*
		(0.111)
Time since founding	-5.92***	-5.91***
1 year	(0.29)	(0.29)
Time since founding	-5.11***	-5.10***
1-2 years	(0.27)	(0.27)
Time since founding	-4.69***	-4.68***
2-5 years	(0.26)	(0.26)
Time since founding	-5.43***	-5.43***
5-10 years	(0.29)	(0.29)
Time since founding	-6.79***	-6.77***
More than 10 years	(0.36)	(0.36)
Firm-level controls	Yes	Yes
Technology field-level controls	Yes	Yes
Industry dummies	Yes	Yes
Year fixed effects	Yes	Yes
Log pseudolikelihood	-2465.59	-2462.47

(a). The Chinese sample (N=6,140)

Variable	Model 7b	Model 8b
Evaluation criteria consensus		0.25*
		(0.10)
Time since founding	-5.30***	-5.27***
1 year	(0.71)	(0.71)
Time since founding	-5.16***	-5.13***
1-2 years	(0.70)	(0.70)
Time since founding	-5.63***	-5.60***
2-5 years	(0.70)	(0.70)
Time since founding	-5.86***	-5.83***
5-10 years	(0.72)	(0.72)
Time since founding	-5.84***	-5.78***
More than 10 years	(0.72)	(0.72)
Firm-level controls	Yes	Yes
Technology field-level controls	Yes	Yes
Industry dummies	Yes	Yes
Year fixed effects	Yes	Yes
Log pseudolikelihood	-1388.17	-1386.15

Note. Control variables and industry dummies are omitted for brevity. Beta coefficients are reported. Standard errors are in parentheses. All two-tailed tests. Robust standard error clustered by firm. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 6. Study 3. Additional Analysis: Effects of Evaluation Criteria Consensus Overtime

(a). The U.S. sample (N=27,837)

Variable	Model 9a
Evaluation criteria consensus	12.71*
	(62.69)
Year	0.02*
	(0.01)
*Evaluation criteria consensus * year*	-12.66*
	(0.03)
Controls	Yes
Industry dummies	Yes
Year fixed effects	No
Log pseudolikelihood	-6616.40

(b). The Chinese sample (N=6,140)

Variable	Model 9b
Evaluation criteria consensus	18.49***
	(65.67)
Year	0.14***
	(0.01)
*Evaluation criteria consensus * year*	-18.46***
	(0.03)
Controls	Yes
Industry dummies	Yes
Year fixed effects	No
Log pseudolikelihood	-5311.85

Note. Control variables and industry dummies are omitted for brevity. Standardized beta coefficients are reported. Standard errors are in parentheses. All two-tailed tests. Robust standard error clustered by firm. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 7. Study 3. Additional Analysis: Contingent Effect of Evaluation Criteria Consensus

The U.S. sample (N=27,837)

Variable	Model 10	Model 11	Model 12
Evaluation criteria consensus	0.04***	0.05*	0.00
	(0.12)	(0.22)	(0.12)
Investor experience	0.07***		
	(0.00)		
*Evaluation criteria consensus * Investor experience*	-0.01**		
	(0.01)		
Professional VC firm		0.25***	
		(0.13)	
*Evaluation criteria consensus * Professional VC firm*		-0.02*	
		(0.22)	
Syndicate size			0.06***
			(0.02)
*Evaluation criteria consensus * Syndicate size*			0.02***
			(0.02)
Controls	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes
Log pseudolikelihood	-6286.32	-5465.96	-6200.30

Note. Control variables and industry dummies are omitted for brevity. Standardized beta coefficients are reported. Standard errors are in parentheses. All two-tailed tests. Robust standard error clustered by firm. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

[i] Per our definition, to have a high level of consensus in evaluation criteria, it is not necessary that experts agree on one dominant criterion; there could be multiple criteria. As long as experts agree on the relative importance of each one and that a competent firm needs to satisfy the multiple criteria simultaneously, there is still high consensus in the field.

[ii] Signals about a firm's technological competence could affect the impact of evaluation criteria consensus. We conducted an additional experiment to investigate the moderating effect about such competence signals. For details about experimental procedures and findings, please refer to Online Appendix G.

[iii] We pre-registered to recruit 60 participants for Study 1. However, when the sample size reached 60, results from power analysis suggested a sample of 80 was needed to get $\alpha = 0.05, 1 - \beta = 0.80$.

[iv] Since 2010, the ImageNet project runs an annual software contest, the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), where software programs compete to classify and detect objects and scenes. AI researcher Fei-Fei Li started the ImageNet project as a large visual database designed for use in visual object recognition software research.

[v] The h5-index for a conference is the largest number h such that h articles published at the conference in the past complete 5 years have at least h citations each.